United States Department of Agriculture

Office of the Chief Information Officer
Network Engineering Division

Telecommunications Enterprise Network Design

Network Level Traffic Study of USDA Data Networks
Task II Report

December 16, 1997
Revised June 9, 1998

Sensitive Information

The section of this document with Agency specific information is considered sensitive. The nature of the information could pose a security risk if it were disclosed to unauthorized personnel. Please consider the sensitivity of this information before making copies and distributing this report.

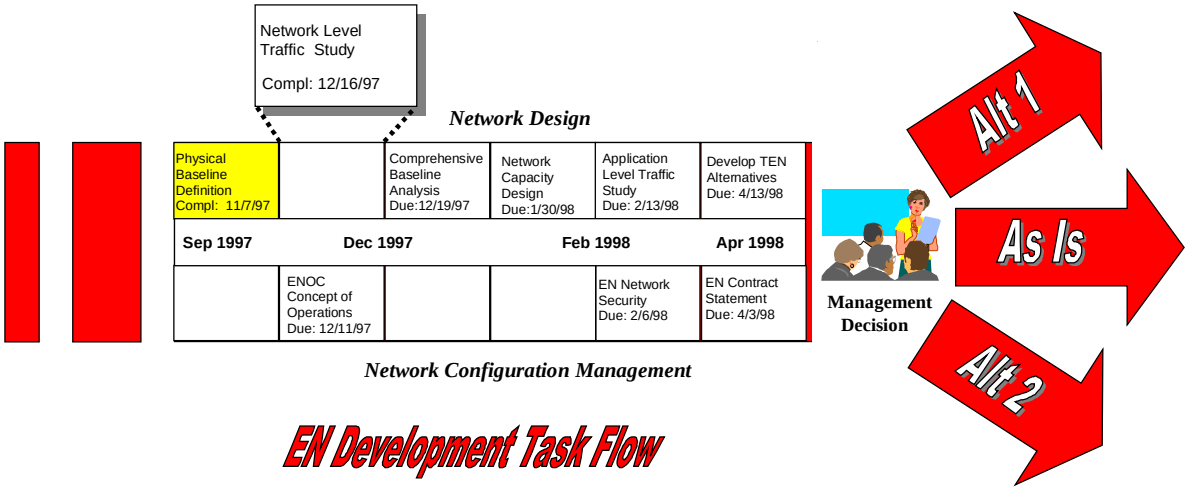
Executive Summary

The US Department of Agriculture, Office of the Chief Information Officer, Network Engineering Division, is responsible for designing and implementing the USDA Telecommunications Enterprise Network according to the Telecommunications Stabilization and Migration Program Plan. The Initial Telecommunications Enterprise Network Design Plan involves six distinct tasks, the second of which is a Network Level Traffic Study of USDA Data Networks. (See EN Development Task Flow Chart below.)

The Task II Report for the Network Level Traffic Study describes the amount of data traffic on the existing network interconnections. Measurement of network data traffic used the method of electronically polling network components and retrieving usage data for each interconnection. The electronically retrieved data is imported into a database and processed.

Summary information is presented regarding the method and frequency of the electronic polling process. Based on the number of circuits identified in the November 7, 1997 Physical Baseline Report, data have been collected for 502 interconnections thus far. The January 30, 1998 discovery identifies 702 interconnections. Data will be collected on the additional circuits as the design process proceeds.

Along with this summary report, each Agency or Organization within USDA receives a detailed list of its interconnections that have been analyzed. For each interconnection, detailed data is summarized and presented in both tabular and graphical format. Utilization analysis is also presented for each interconnection.



1.0 Introduction

The heart of the Telecommunications Network Stabilization and Migration Program (TSNMP) is network optimization – maximizing performance while reducing operating costs. Key to the optimization theme is standardization of equipment and processes. To accomplish network uniformity, the Geographic Network Analysis Process (GNAP) has been developed. This process promotes objective and uniform assessment of network operation and the initiation of improvements. The GNAP, Figure 1, is used by the Network Engineering Division to evaluate the existing USDA Data Networks and create an optimized Initial Telecommunications Enterprise Network (TEN).

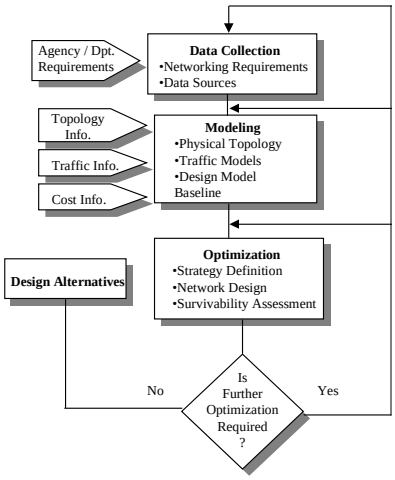


Figure 1 Geographic Network Analysis Process

The USDA IEN Design Plan is divided into six distinct tasks:

- Physical Baseline Definition of the USDA Data Networks
- Network Level Traffic Study of the USDA Data Networks
- Comprehensive Baseline Analysis of USDA Data Networks
- Development of Network Capacity Design
- Application Level Traffic Study of the USDA Data Networks
- Development of Initial TEN Design

A prerequisite to the development of the comprehensive TEN design is a complete and accurate description of the data networks existing throughout the USDA. The first three tasks of the USDA Initial TEN Design Plan represent a data collection phase, the combined results of which describe the current state of the network – the Network Baseline. The Network Baseline is a dynamic model, refined as new information is derived or made available. The model represents a "snapshot" of the Network Baseline at any particular time prior to migration to the Initial TEN.

This report addresses the second task of the data collection phase, the Network Level Traffic Study of USDA Data Networks. The goal of the Network Level Traffic Study is a comprehensive description of the utilization of Wide Area Network links by all USDA networks.

Due to time constraints for inventory assessment and the need to provide accurate information, this report addresses only a subset of the network links. As additional information on the physical topology of the networks becomes available, the Network Level Traffic Study is updated accordingly.

Goals of this report are both to inform and learn. The information presented is to inform Agencies about the knowledge of existing USDA Networks. The accuracy and completeness of the USDA Network description is enhanced by Agency feedback correcting erroneous information and adds missing information.

2.0 Methodology: Network Level Traffic Data Collection

Network level Traffic data is collected using electronic inquiry of known network equipment. The NetMaker®[1] XA Simple Network Management Protocol (SNMP) Agent dynamically polls routers to obtain the information stored inside each router's Management Information Base (MIB). This information includes usage statistics for each interface on the router.

The network model used in the data collection is based on a previous automated discovery. Circuit speeds for the usage analysis are taken from the router's configuration. Based on this information, NetMaker® XA automatically calculates the minimum number of routers that need to be polled in order to collect information for all existing circuits. The software then polls these designated routers during the specified sampling period and collects the usage data.

Once the raw usage data is collected, profiles for high and average usage are created for each WAN Link per sampled time slice. This data is then imported into the NetMaker® XA modeling tool and analyzed using the network model derived from the automated discovery process. The resulting usage analysis, along with the raw average and high usage data, is presented in tabular and graphical reports.

3.0 Results

Initial results are presented in two parts. The network level usage analysis for all of USDA is presented in section 3.1. The addendum of this report details results for the respective Agency or Organization. Addendum A contains raw average and high usage data for each sampled link. Addendum B contains usage analysis for each link.

[1] NetMaker® XA is a registered trademark of Make Systems, Inc.

The routers used in the data collection process are taken from the list of routers designated by the automated discovery dated December 1, 1997 and January 14, 1998. Samples were collected from 6 A.M. EST to 6 P.M. PST on weekdays, on thirteen days between January 12, 1998 and February 6, 1998. Holidays and the Friday after Thanksgiving were excluded. The data collected is organized into 16, one hour samples for each day. For each time slice, the samples collected represent the number of bytes and the number of packets transferred in and out for each link.

The sample data is organized to represent an "average" and a "high" usage day. The "average" day is calculated as the average of each one hour time slice for all of the sampled days. The "high" day is calculated as the high value of each one hour time slice for all of the sampled days. The "high" day is used in the worst case analysis for the sample time period. These two calculated usage days are superimposed on the network model produced from a network discovery dated January 30, 1998 to describe network level traffic.

The more recent discovery is used to include router configuration changes made as a result of information presented in the Physical Baseline Report[2]. This additional data makes circuit bandwidth values more accurate as well as introducing circuits that had not previously been discovered. Since these circuits are not included in the November 29, 1997 discovery, there is no usage data available. They are included in this report to complete the circuit inventory.

It should also be noted that Frame Relay Permanent Virtual Circuits (PVC) are not included in the analysis. Information on the Frame Relay PVC is currently not obtainable using the SNMP engine and router MIB. The only portion of the Frame Relay network that is analyzed is the Frame Relay access circuit between the router and the Frame Relay Point-of-Presence (POP). Even though this limitation affects the network level traffic study, it does not impede the development of IEN design. The remainder of the design process is not dependent on the usage statistics obtained from the router.

3.1 Network Level Usage Summary

Since usage data sampling is based on the November 7, 1997 Baseline Report, there is only data available for 502 circuits. To improve the accuracy of the line speeds reported in router configurations, the January 30, 1998 discovery is used in the usage analysis. This discovery identified a total of 702 circuits. As a result 175 circuits appear in the analysis as zero utilization. This discrepancy skews average data such that values are artificially low. It is important to note that the average results validate methodology and in no way affect the accuracy of the data presented in Addendum A and B.

[2] Telecommunications Network Design, Physical Baseline Definition of USDA Data Networks, Task I Report. November 7, 1997. Network Engineering Division

As usage data is collected for the additional circuits and line speed information is corrected in router configurations, the accuracy of the network level traffic analysis improves. The data summary presented demonstrates the type of information available from the NetMaker® XA modeling tool.

"Average" Day Analysis

Average WAN Link Utilization: 4.037 %
Minimum WAN Link Utilization: 0.000 %
High WAN Link Utilization [3]: 131.653%

"High" Day Analysis:

Average WAN Link Utilization: 19.116 %
Minimum WAN Link Utilization: 0.000 %
High WAN Link Utilization[3]: 244.311 %

Table 1 summarizes the initial inventory of USDA WAN Links as of January 30, 1998. It should be noted that this inventory is based on router configuration information obtained during the January 30, 1998 automated discovery.

Number of WAN Links by type or speed	
0-64K	**225**
65-128K	**61**
129-256K	**52**
385-512K	**17**
513-768K	**35**
769-1024K	**4**
1025-1536K	**49**
T1	**258**
E1	**1**
Total	***702***

Table 1 Number of Wide Area Network Links

4.0 Network Level Traffic Data Caveats

- The accuracy of the numbers for each link speed is determined by the accuracy of router configurations. Most routers use a default speed of 1.544Mbps (T1) if the configuration is either incomplete or inaccurate. The inordinately large number of reported T1 WAN links (Table 1) reflects the scope of the problem. As router

[3] The percentages for the high utilization are greater than 100 because of incorrect (low) line speeds reported in some router hardware. These line speeds are correctable in the router configuration as discussed in Section 5 of this document.

configurations are corrected, many links reported as T1 are placed in proper categories.

- The accuracy and completeness of the data presented in this report for the USDA as a whole, is limited by the completeness and accuracy of the network model presented in the physical baseline report.
- Network Level Traffic analysis is based on data from two different discoveries causing reported traffic usage values to be skewed toward the low end.
- The results presented in this report use the process described in section 2.0. The results validate the methodology chosen for the data collection phase of the project.

5.0 Discussion/Recommendations

This report represents analysis of USDA network utilization based on the best data available and validates the methodology chosen for the data collection phase of the Initial TEN Design Plan. The network level traffic study will continue to be updated as new routers and circuits are discovered, and as information on known circuits is updated in router configurations. As the accuracy and completeness of the network model is improved, the accuracy of the network level usage analysis will improve.

The network level traffic data presented in this report, and subsequent data collection, is used to provide objective information on which to base recommendations to change the capacity of existing network links. Such recommendations are the subject of the Initial TEN Design Plan deliverable entitled "Network Capacity Design – Task IV." Recommendations to change link size are based on utilization analysis and the Agency or application performance requirements. This means that an apparent low use link should not be reduced if the application requires the throughput for a given transaction. All known factors will be used to develop recommendations.

The data for this report, collected in November and early December, reflects historically low utilization months for the USDA. The analysis of network capacity takes this information into consideration when identifying those circuits that need to be considered for change. Additional usage data to augment the data presented in this report is to be collected in early January.

The data presented in the addendum includes specific Agency circuits discovered after the Physical Baseline Report and not used in the Network Level Traffic Study. However, these circuits are to be included in future usage data collections and are included in this report so that the circuit speeds may be verified. Discrepancies in the circuit speeds presented in the addendum should be corrected in the corresponding router configuration as described in the following paragraphs.

Verification of Report information: The information provided in the addendum of this report represents the current description of the respective Agency telecommunications

network. It is extremely important that time be allocated to check and make appropriate changes to this information in the appropriate router configurations.
It is of particular importance to examine the line speeds reported for each of the links. This speed is used in the ongoing analysis of network utilization and cost. If a discrepancy is found, it must be corrected in the router's configuration. The following router command is used to set/correct the line speed used in the automated analysis. This command is for the Cisco Internetworking Operating System (IOS), which is the predominant router software on the USDA network. For non-Cisco IOS routers, the equivalent configuration for the device should be approximated.

Interface "bandwidth" Command:

This command is used to set the SNMP variable used to determine the bandwidth of the circuit connected to each router interface. If this command is not used, the router reports the maximum bandwidth capability of the router interface hardware. For example, if a 56Kbps circuit is connected to a router interface that is capable of supporting a full T1, and the bandwidth command is missing, the circuit on the interface will be modeled as a full T1. This discrepancy will skew the circuit utilization and cost analysis. It is also important that interfaces on each end of a circuit have matching bandwidth values.

The following is an example of how to use the bandwidth command to set the circuit line speed of a router interface to 56Kbps.

```
Interface Serial0
bandwidth 56
```

NB The bandwidth number is in units of Kbps. Do not append trailing zeros.

FOR HELP CONTACT THE NETWORK DESIGN TEAM:

Dan Crosson
OCIO Telecommunications Services, Network Engineering Division
2627 Redwing Road
Creekside 1, Second Floor
Fort Collins, CO 80526
E-Mail: dcrosson@tasc.usda.com
Voice: (970) 282-1909
FAX: (970) 282-1999

Addendum A – WAN Link Traffic Report

This addendum provides summary information on network link traffic samples. A network link is defined as any connection between two routers. The link may be Dedicated Transmission Service, Frame Relay, or LAN based. This summary does not make distinctions based on the make up of the link between the two routers, but only reports the amount of traffic that was passed between the routers.

The sample data are organized to represent an "average" and a "high" usage day. The "average" day is calculated as the average of each one hour time slice for all of the sampled days. The "high" day is calculated as the high value of each one hour time slice for all of the sampled days. The following should be noted:

- The "Source" and "Destination" fields contain router names associated with the link end points.

- The columns in the table are defined as follows.

  Sample Time - The start of the fifteen-minute time slice for the data totals included in the row.

  No. of Samples - The number of data samples used for calculating the average and high values presented in the row.

  Averages (Forward) – The average number of bytes and packets transmitted from the source router for the time slice presented in the row.

  Averages (Return) – The average number of bytes and packets received by the source router for the time slice presented in the row.

  Highest Samples (Forward) – The number of bytes and packets transmitted from the source router corresponding to the time slice sample containing the highest total number (forward and return) of bytes transmitted.

  Highest Samples (Return) – The number of bytes and packets received by the source router corresponding to the time slice sample containing the highest total number (forward and return) of bytes transmitted.

Sensitive Information
Disclosure of information in this section to unauthorized personnel could pose a security risk.

Addendum B – Wide Area Network Links Usage Analysis

This addendum provides detailed usage information on the wide area network links connected to router hardware. The numbers presented in this report are calculated based on the network model from the December 1, 1997, automated discovery. The data presented in Addendum A was then imported into the discovery and analyzed using the network modeling tools.

The "Line Speed" field for these links should be verified. Any discrepancies or missing information should be corrected in the appropriate router configuration and reported to the Network Design Team. The following should be noted:

- The "Line Speed" field is taken from router MIB configurations and must be set correctly, or the circuit's utilization is skewed. Section 5.0 of this document describes how to set the circuit's "Line Speed".

- The order in which the end point nodes are presented was determined by an automated discovery process, and does not imply any hierarchical relationship.

- The "Node A" and "Node B" fields contain the router names of the link end points.

- The "Source" and "Destination" fields contain router names in the case of point-to-point circuits. When one end of a link is connected to a Frame Relay POP, one of the fields contains a "P" with the router's name appended.

- The line speed used in the usage analysis is based on information from a December 1, 1997 discovery of the network. The "Average" and "High" days used in the analysis are taken from the data presented in the Addendum A table.

Sensitive Information
Disclosure of information in this section to unauthorized personnel could pose a security risk.